Introduction

An overriding goal of outcomes research is measuring and comparing hospital performance.

Hospital profiling is a hot topic because provider-specific outcomes, such as mortality rates, are used to:
- Designate centers of excellence
- Determine reimbursement levels in pay for performance programs
- Classify providers as outliers
Risk-adjustment (1)

The outcome rates need to be adjusted to account for pre-existing factors that are outside the quality of care delivered and may confound their assessment.

We all know the direct and indirect methods to calculate age-adjusted rates.

Things get more complicated when other variables, such as clinical factors, are taken into account to derive risk-adjusted rates. These additional characteristics can be retrieved from the hospital discharge abstracts, which are generally inexpensive and enable the analysis of large populations and a large number of clinical conditions.
Risk-adjustment (2)

Statistical techniques to derive hospital-specific risk-adjusted rates can be trivially outlined into three steps:

- Fit an appropriate regression model to the data
- Calculate risk-standardized outcomes resulting from regression analysis
- Display risk-standardized outcomes using appropriate graphical representations

(We are not going to review advanced techniques such as Bayesian hierarchical models.)
The easiest way to obtain risk-adjusted outcome measures across providers is to build a “conventional” logistic regression model where \textit{depvar} is a binary outcome measure expressed as 0/1 (say death) and covariates \textit{indepvars} are the patient case mix (age, sex, co-morbidities).

\texttt{logit depvar indepvars [if] [in]}

However, some precautions need to be taken to account for the positive correlation between observations from within the same hospital.
Robust standard errors

When conducting analyses on correlated data, a robust estimation of standard errors should be used. These are known as robust, “sandwich”, or Huber-White standard errors. In this command, `varname_i` uniquely identifies providers.

```
logit depvar indepvars [if] [in], vce(cluster varname_i)
```

These estimates allow the correct specification of the mean model while relaxing the assumption of correctly specifying the form of the variance model, i.e., the working correlation matrix.
Generalized estimating equations (GEEs)

Fitting conventional regression models to correlated data might also lead to inefficient parameter estimates.

GEEs, a flexible extension of conventional regression models, account for correlated observations. In Stata, the default GEE within-group correlation structure corresponds to the equal-correlation model, also called “exchangeable”, which is appropriate for profiling studies (Ballinger 2004).

```
xtgee depvar indepvars [if] [in], family(binomial) link(logit) i(varname_i) vce(robust) corr(exchangeable) // c(exc) is optional
```

Robust standard errors should be estimated in conjunction with GEEs (Liang and Zeger 1986) by using the vce(robust) option.
Conventional models or GEEs?

Results from GEEs and logistic regression are identical if the within-group correlation is close to zero.

To test whether observations are actually correlated, one should compare a GEE model with an “exchangeable” working matrix and a GEE model with an “independent” working matrix, the latter corresponding to logit. The best model has the lowest quasi-likelihood under independence criterion (Pan 2001).

The `qic` command is downloadable from SSC (Cui 2007).

```
qic depvar indepvars [if] [in], family(binomial) link(logit) i(varname_i) robust corr(exchangeable)
qic depvar indepvars [if] [in], family(binomial) link(logit) i(varname_i) robust corr(independent)
```
Predicted probabilities

After running the best model between the two, we estimate individual risk for each patient (newvar) using the observed values of her/his confounding variables.

predict newvar [if] [in], rate // , pr in logit post-estimation

Then the predicted probabilities can be summed over all records within each provider to derive the expected number of events.

tabstat newvar [if] [in], by(varname_i) statistic(sum)

The expected number of events is the number of events that would occur if the “standard” event rates had happened, given the provider actual case mix. This can be seen as an “indirect” form of standardization.
Risk-standardized rates

The adjusted outcome for each provider is presented as the ratio of the observed to the expected number of events (O/E ratio). The O/E ratio is favorable if <1, and unfavorable if >1.

To allow for comparison of each hospital performance with the national/regional average, the O/E ratio should be multiplied by the overall outcome rate. This measure is called risk-standardized rate. When the events are deaths, the common name is risk-standardized mortality rate (RSMR).

Before discussing how to get these measures in Stata, let us introduce an example based on real data.
Example: 30-day AMI mortality rates in Latvia

We use real data from 20 hospitals in Latvia, year 2016. A total of 2916 patients with acute myocardial infarction (AMI) met the inclusion criteria, and the overall mortality rate was 17.5%.

Death within 30 days of hospital admission for AMI is \( \text{Death30Days} \), expressed as 0/1, and the hospital identification number is \( \text{HospitalID} \). Potential confounders include: age, sex, ST elevation status, history of AMI, and 31 co-morbidities based on the Elixhauser method (Quan et al. 2005).

Variables included as confounders in the multivariable model from which RSMRs will be computed are: age, sex, cardiac arrhythmias, pulmonary circulation disorders, peripheral vascular disease, diabetes with complications, and ST elevation status.

... Wait a minute: why were these specific variables selected for inclusion?
Confounder selection

The choice of predictive variables in regression analysis is somewhat of an art. When we use pre-defined sets of covariates, such as Elixhauser’s, we may want to avoid model over-fitting and misclassification by including only covariates whose estimated regression coefficients are stable. This can be achieved via numerous automated selection procedures.

See Austin and Tu (2004) for methodological details, and Royston and Sauerbrei (2009) to see how these automated selections can be performed in Stata.
Choosing the right model (1)

We first calculate the QIC value for the exchangeable correlation structure, then we do the same for the independent correlation structure.

```
.sset matsize 11000 // default matrix size must be augmented first
.xi: qic Death30Day Sex i.AgeCL4 CARDARRH PULMCIRC PERIVASC DMCX i.AMItpe, ///
  > family(binomial) link(logit) i(HospitalID) corr(exchangeable) ///
  > nolog nodisplay // GEE model with nolog and nodis options to save space

| i.AgeCL4 | _IAgeCL4_1-4          (naturally coded; _IAgeCL4_1 omitted) |
| i.AMItpe | _IAMItpe_1-3          (naturally coded; _IAMItpe_1 omitted) |
```

QIC and QIC_u

```
-------------------------------------------
Corr = exchangeable
Family = binomial
Link = logit
p = 11
Trace = 13.999
QIC = 2479.168
QIC_u = 2473.171
-------------------------------------------
```
Choosing the right model (2)

```
.xi: qic Death30Day Sex i.AgeCL4 CARDARRH PULMCIRC PERIVASC DMCX i.AMItype, ///
> family(binomial) link(logit) i(HospitalID) corr(independent) ///
> nolog nodisplay // logistic model with nolog and nodis options to save space
i.AgeCL4          _IAgeCL4_1-4        (naturally coded; _IAgeCL4_1 omitted)
i.AMItype _IAMItype_1-3       (naturally coded; _IAMItype_1 omitted)
```

QIC and QIC_u

-------------------------------------------
Corr =          independent
Family =           binomial
Link =                logit
p =                      11
Trace =              18.931
QIC =              2429.175
QIC_u =            2413.313
-------------------------------------------

The exchangeable structure has a QIC of 2479.168, while the independent structure has a QIC of 2429.175.
Choosing the right model (3)

Conventional logistic regression is the best fitting model here.

```
.xi: logit Death30Day Sex i.AgeCL4 CARDARRH PULMCIRC PERIVASC DM CX i.AM Itype, or vce(cluster HospitalID) nolog
i.AgeCL4          _IAgeCL4_1-4        (naturally coded; _IAgeCL4_1 omitted)
i.AM Itype _IAMItype_1-3       (naturally coded; _IAMItype_1 omitted)
```

Logistic regression

```
Logistic regression                              Number of obs =      2,916
Wald chi2(10)        =    1388.99
Prob > chi2          =     0.0000
Log pseudolikelihood = -1195.6566               Pseudo R2         =     0.1155
```

(Std. Err. adjusted for 20 clusters in HospitalID)

<table>
<thead>
<tr>
<th></th>
<th>Robust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>Death30Days</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>1.046859</td>
</tr>
<tr>
<td>_IAgeCL4_2</td>
<td>2.342332</td>
</tr>
<tr>
<td>_IAgeCL4_3</td>
<td>3.72511</td>
</tr>
<tr>
<td>_IAgeCL4_4</td>
<td>7.319825</td>
</tr>
<tr>
<td>CARDARRH</td>
<td>1.835576</td>
</tr>
<tr>
<td>PULMCIRC</td>
<td>2.372894</td>
</tr>
<tr>
<td>PERIVASC</td>
<td>1.802321</td>
</tr>
<tr>
<td>DM CX</td>
<td>2.315345</td>
</tr>
<tr>
<td>_IAMItype_2</td>
<td>.4575801</td>
</tr>
<tr>
<td>_IAMItype_3</td>
<td>1.312559</td>
</tr>
<tr>
<td>_cons</td>
<td>0.0513947</td>
</tr>
</tbody>
</table>
```

Note: _cons estimates baseline odds.
Hospital-specific RSMRs following AMI (1)

After logit post-estimation ($p_{\text{hat}}$), a few command lines have to be run to derive hospital-specific RSMRs and save these values as a new data file.

First, use `collapse` to manipulate data at hospital level and to derive the total number of observed deaths ($\text{Obs}$), expected deaths ($\text{Exp}$), and patients ($\text{N}$)—remember to launch `preserve` first! For each hospital, you generate a new variable ($\text{MR}$) containing the crude mortality rates (%).

```stata
.predict $p_{\text{hat}}$, pr
.preserve
.collapse (sum) Obs = Death30Day Exp = $p_{\text{hat}}$ (count) N = Death30Day, /// > by(HospitalID)
.generate MR = Obs*100/N
```
Hospital-specific RSMRs following AMI (2)

After `tabstat`, we define a scalar (`Rate`) containing the overall mortality rate value, that is then used to calculate the RSMRs. We also compute the 95% CIs for RSMRs based on the formula that relates the chi-squared distribution and the Poisson distribution (Garwood 1936).

```
. tabstat Obs N, statistic(sum) save

     stats  |     Obs     N
      +------------------+
       sum  |       510     2916
      +------------------+

. matrix total = r(StatTotal)

. scalar Rate = total[1,1]*100/total[1,2]

. generate RSMR = Obs/Exp*Rate

. generate lb_RSMR = (invchi2(2*Obs,0.05/2)/2)/Exp*Rate // RSMR lower limit
    (1 missing value generated)

. generate ub_RSMR = (invchi2(2*Obs+2,1-0.05/2)/2)/Exp*Rate // RSMR upper limit
```
Hospital-specific RSMRs following AMI (3)

New results are saved as `RSMR.dta` and visualized with `list`.

```
. save RSMR
. list

<table>
<thead>
<tr>
<th>HospitalID</th>
<th>Obs</th>
<th>Exp</th>
<th>N</th>
<th>MR</th>
<th>RSMR</th>
<th>lb_RSMR</th>
<th>ub_RSMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>3.260993</td>
<td>23</td>
<td>26.08696</td>
<td>32.17986</td>
<td>11.80944</td>
<td>70.042</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>6.318017</td>
<td>38</td>
<td>13.1579</td>
<td>13.84114</td>
<td>4.494181</td>
<td>32.30061</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>9.483574</td>
<td>46</td>
<td>30.43478</td>
<td>25.81895</td>
<td>14.11546</td>
<td>43.31982</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
<td>35.05648</td>
<td>234</td>
<td>15.38461</td>
<td>17.96043</td>
<td>12.57927</td>
<td>24.86481</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>4.12641</td>
<td>21</td>
<td>14.28571</td>
<td>12.71544</td>
<td>2.62223</td>
<td>37.15992</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>20.86314</td>
<td>102</td>
<td>22.54902</td>
<td>19.28106</td>
<td>12.22254</td>
<td>28.93106</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>17.96316</td>
<td>116</td>
<td>24.13793</td>
<td>27.26201</td>
<td>18.11541</td>
<td>39.4012</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>6.220828</td>
<td>35</td>
<td>34.28571</td>
<td>33.73772</td>
<td>17.43277</td>
<td>58.933</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>2.108751</td>
<td>7</td>
<td>57.14286</td>
<td>33.1755</td>
<td>9.039207</td>
<td>84.94244</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>4.146946</td>
<td>26</td>
<td>23.07692</td>
<td>25.30495</td>
<td>9.286472</td>
<td>55.07823</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>7.993699</td>
<td>52</td>
<td>7.692307</td>
<td>8.75175</td>
<td>2.384557</td>
<td>22.40796</td>
</tr>
<tr>
<td>13</td>
<td>14</td>
<td>12.1716</td>
<td>60</td>
<td>23.33333</td>
<td>20.11699</td>
<td>10.99815</td>
<td>33.75289</td>
</tr>
<tr>
<td>14</td>
<td>91</td>
<td>124.8035</td>
<td>709</td>
<td>12.83498</td>
<td>12.75255</td>
<td>10.26756</td>
<td>15.65731</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>.3216332</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>142</td>
<td>152.1697</td>
<td>849</td>
<td>16.72556</td>
<td>16.32086</td>
<td>13.74692</td>
<td>19.23678</td>
</tr>
<tr>
<td>18</td>
<td>17</td>
<td>7.870675</td>
<td>36</td>
<td>47.22222</td>
<td>37.77632</td>
<td>22.0061</td>
<td>60.48354</td>
</tr>
</tbody>
</table>
```

. restore // restore original data after preserve
Graphical representations of RSMRs

RSMR.dta is now used to produce plots of hospital-specific RSMRs.

Plots useful to display outcome rates are (Spiegelhalter 2005):
- Caterpillar plots, sometimes referred to as “forest plots”
- Funnel plots

Both plots, especially funnel plots, operate a fair distinction between small and large providers, and allow to ascertain whether large deviations from the state average are systematic or due to chance.
The caterpillar plot is a sort of league table where providers are ranked according to a performance indicator and, with the aid of CIs, outlying providers are identified. To avoid data misinterpretation, the providers should never be labeled with their rank, and outlying positions have to be strictly determined using CIs.

Plots of estimates and CIs can be obtained in Stata using `eclplot`, which is downloadable from SSC (Newson 2003).
Caterpillar plot of RSMRs following AMI (1)

Before launching `eclplot`, a new variable with the ranking of hospitals (`Rank`) must be created. Hospital #15, with only one AMI patient, has been previously discarded.

```
. sort RSMR // Sorting hospitals according to RSMRs

. generate Rank = _n

. eclplot RSMR lb_RSMR ub_RSMR Rank, plotregion(color(white) ilcolor(black) ///
> margin(none)) graphregion(color(white)) ylabel(0(10)100, angle(360) notick ///
> nogrid) yscale(noextend noline) ytitle("RSMR (%)") xlabel(0 " " 20 " ", ///
> notick) xscale(noextend noline) xtitle("")) estopts(mlabel(HospitalID) ///
> mlabposition(0) mlabcolor(white) msymbol(o) msize(huge) mcolor(gs6)) ///
> ciopts(msize(zero) lwidth(medthick) lcolor(gs10)) ///
> baddplot(function y = Rate, col(black) lwidth(thin) range(0 20))
```
Caterpillar plot of RSMRs following AMI (2)

Hospital #14 is a low outlier, while hospitals #7 and #18 are high outliers.
Funnel plot

Each hospital-specific RSMR (y axis) is plotted relative to its denominator size (x axis). Given \( r \) as the overall rate, \( n \) as the hospital volume and \( z \) as the standard normal distribution quantile, control limits are superimposed following the formula:

\[
y_{\alpha/2}(r, n) = r \pm z_{\alpha/2} \sqrt{\frac{r(1-r)}{n}}
\]

where \( z_{\alpha/2} \) is 1.96 for 95% control limits and 3.09 for 99.8% control limits.

These boundaries are a measure of precision of the hospital rates, and depend on denominator sizes. Hospitals lying outside the control limits can be seen as outliers.
Funnel plot of RSMRs following AMI (1)

Funnel plots can be obtained either using the {	t funnelcompar} package (Fomi and Gini 2009), or combining a scatter plot with two-way function plots.

We have opted for the second choice. Before doing that, we set highest and lowest values for control limits, based on the range of $y$ axis (up to 60% here).

\begin{verbatim}
. local up95 = (Rate*(100-Rate))/((60-Rate)/1.96)^2 // Highest 95% limit point
. local low95 = (Rate*(100-Rate))/((0-Rate)/1.96)^2 // Lowest 95% limit point
. local up99 = (Rate*(100-Rate))/((60-Rate)/3.09)^2 // Highest 99.8% limit point
. local low99 = (Rate*(100-Rate))/((0-Rate)/3.09)^2 // Lowest 99.8% limit point
\end{verbatim}
Funnel plot of RSMRs following AMI (2)

The Stata syntax for a customized funnel plot with the range of $x$ axis up to 900 is as follows.

```
. twoway ///
  > (function y = Rate+1.96*sqrt((Rate*(100-Rate))/x), col(black) lwidth(thin) ///
  > lpattern(dash) range(`up95' 900)) ///
  > (function y = Rate-1.96*sqrt((Rate*(100-Rate))/x), col(black) lwidth(thin) ///
  > lpattern(dash) range(`low95' 900)) ///
  > (function y = Rate+3.09*sqrt((Rate*(100-Rate))/x), col(black) lwidth(thin) ///
  > lpattern(shortdash) range(`up99' 900)) ///
  > (function y = Rate-3.09*sqrt((Rate*(100-Rate))/x), col(black) lwidth(thin) ///
  > lpattern(shortdash) range(`low99' 900)) ///
  > (function y = Rate, col(black) lwidth(thin) range(0 900)) ///
  > (scatter RSMR N, plotregion(color(white) ilcolor(black) ///
  > margin(none)) ytitle("RSMR (%)") xtitle("AMI Patients", height(5)) ///
  > ylabel(-0(10)60, angle(360) glcolor(gs15) glwidth(vthin) nogmax nogmin notick) ///
  > xlabel(0(100)900, grid glcolor(gs15) glwidth(vthin) nogmax nogmin notick) ///
  > xscale(noextend noline) yscale(noextend noline) mlabcolor(black) ///
  > mlabsize(vsmall) mfcolor(none) mlcolor(black) mlwidth(thin) bgcolor(white) ///
  > graphregion(color(white)) mlabel(HospitalID) mlabposition(0) msymbol(o) ///
  > msize(huge) mlcolor(gs8) mfcolor(gs8) mlabcolor(white) ///
  > legend(order(3 "99.8% Control Limit" 1 "95% Control Limit" 5 "Overall Mortality Rate") ///
  > size(small) col(1) ring(0) pos(2) region(style(none))))
```
Funnel plot of RSMRs following AMI (3)

Hospital #14 is a low outlier, while hospitals #7, #8 and #18 are high outliers.

RSMRs following AMI in 19 hospitals in Latvia, Year 2016. The target is the overall rate of 17.5%. Hospital #15 is excluded.
Caterpillar plots or funnel plots?

Funnel plots might be preferred to caterpillar plots because:
- The eye is naturally drawn to important points that lie outside the funnel
- There is no spurious ranking of institutions
- The relationship of outcome with volume can be preliminarily assessed
- Pairwise comparisons between providers are naturally discouraged
For more details...

... check the next issues of the Stata Journal: “Tips for Calculating and Displaying Risk Standardized Hospital Outcomes in Stata” by J. Lenzi and S. Pildava.
References


Jacopo Lenzi

Dipartimento di Scienze Biomediche e Neuromotorie
Sezione di Igiene generale ed applicata e Statistica medica

jacopo.lenzi2@unibo.it

www.unibo.it